Genomics and bioinformatics enhance research on food and gut bacteria - Editorial overview

Kuipers, OP

Published in:
Current Opinion in Biotechnology

DOI:
10.1016/j.copbio.2004.03.003

ELSEVIER

Genomics and bioinformatics enhance research on food and gut bacteria

Editorial overview

Oscar P Kuipers

Current Opinion in Biotechnology 2004, **15**:83–85

0958-1669/$ – see front matter

DOI 10.1016/j.copbio.2004.03.003

Oscar P Kuipers

Molecular Genetics Group, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Kerklaan 30, 9751 NN Haren, The Netherlands
e-mail: o.p.kuipers@biol.rug.nl

Oscar Kuipers is Professor in molecular genetics of prokaryotes at the University of Groningen. He is involved in functional and comparative genomics studies on a variety of Gram-positive bacteria. His laboratory is applying molecular biology tools and DNA microarrays to elucidate, characterize and visualize complex gene regulatory networks operating in, for example, stress response, carbon and nitrogen metabolism, antimicrobial resistance, pathogenesis, competence development, sporulation and protein secretion. Novel bioinformatics tools are being developed to support the (post)genomics activities.

Abbreviations
LAB lactic acid bacteria
(R-)IVET (recombination-based) *in vivo* expression technology

Food and gastrointestinal tract bacteria have been subjected to thorough studies for many years, because of their importance for fermented food products and, more generally, for human health. But, no other technological advances in the past 40 years have impacted this area so much as the rapid developments in genomics, transcriptomics, proteomics, metabolomics and bioinformatics. Many examples of the use of these technologies, in combination with sound molecular biology and biochemistry, in medicine, pharmacology, and molecular biology are available. These examples illustrate the tremendous power, in particular of the integrated use of these approaches, to elucidate underlying molecular mechanisms of important cellular processes and to discover novel factors involved in important diseases, the pathogenicity of microbes, resistance to antibiotics, and so on.

To date, research on food and gut microbes has also greatly benefited from these high-throughput toolkits and approaches. For many years, biologists interested in engineering bacteria for improved food functionality, safety or flavor, had to go to the lengths of finding the genes involved in these functions by tedious identification methods and through the reconstruction of pathways. Nowadays, the availability of over 200 full genomic microbial DNA sequences (e.g. see http://www.tigr.org/tigr-scripts/CMR2/CMRHomePage.spl) enables a comparative genomics approach for which various novel bioinformatics tools and database systems have been developed. This includes the (partial) construction of metabolic pathways for the organism of interest, superimposing gene regulatory information, the automatic discovery of important regulatory *cis* elements in intergenic regions, the discovery of protein function using a wide variety of tools (e.g. domain, operon, next neighbour and metabolic pathway analyses), and linking this information to the available literature. These efforts have already yielded valuable databases and network information for bacteria such as *Bacillus subtilis*, *Lactococcus lactis*, *Lactobacillus plantarum* and various pathogens.

Even more importantly, many efforts have now been started in the use of experimental high-throughput approaches that enable transcriptome, proteome, interactome and metabolome determinations. This adds a dynamic level to the description of food and gut bacteria, as these are known to vary gene expression, protein synthesis and metabolite production significantly and rapidly during growth, and even more in response to changing growth conditions (e.g. changes in nutrients, pH, temperature, oxygen etc). Concomitantly, specific bioinformatics tools are being developed within and

outside the bacterial research community to store and interpret the wealth of information coming from these time-resolved studies.

In this issue, four papers illustrate advances made in the genomic analysis of food and gut bacteria. Important issues, such as the different functionalities of bacteria in the human gastro-intestinal tract, the impact of bacteriophages on fermented food production and as biotherapeutics, the role of antimicrobial compounds in food safety and last, but not least, the important unifying role of bioinformatics in handling, interpretation and visualization of data, are addressed. These serve as clear examples of how the field of research on food and gastro-intestinal tract bacteria is rapidly moving into the application of integrated functional genomics to facilitate metabolic engineering, optimize *in situ* activities, reduce food spoilage and enhance the development of functional foods.

de Vos, Bron and Kleerebezem discuss various aspects of gut microbe ecology and functionality. They explain how genomics approaches can be used to develop better probiotic cultures and other products that affect gut health. Probiotic cultures are assumed to have a beneficial effect on gut and overall health, and many studies are now being directed towards answering questions about the molecular mechanisms by which these bacteria exert their effects. Increased insight into the genes involved in these processes will further enhance the development of 'health impact' foods and, perhaps even more importantly, will help to support health claims with sound experimental evidence. The gastro-intestinal tract, a highly dynamic organ, harbors more than a thousand different species, many of which cannot easily be cultured. Recently, an initiative has been started to investigate the whole 'microbiome' (i.e. to determine all of the genome sequences of this complex microbial community). How this approach will take into account the considerable differences between individuals has not yet been discussed.

Genome mining of some known 'good' and 'bad' species, like *L. plantarum*, *Enterococcus faecalis and Listeria monocytogenes*, reveals a strikingly high abundance of sugar transport and utilization pathways. In *L. plantarum* these genes are located in so-called 'life style adaptation' islands on its genome. These sugar-utilizing proteins are not only relevant for supporting bacterial growth, but also impact upon the local environmental conditions of colon cells. One of the most interesting aspects of gut bacteria is their ability to interact with host cells and to thrive in the host environment. Therefore, novel methodologies are emerging that enable these properties to be studied, and specifically identify those genes that play a prominent role. One example is the use of DNA microarrays of relevant species such as *L. plantarum* to unravel gene regulatory networks that operate while residing in the host. The authors discuss three main strategies for the discovery of highly expressed genes *in vivo*: (recombination-based) *in vivo* expression technology ((R-)IVET), signature-tagged mutagenesis and selective capture of transcribed sequences. Various important factors were discovered using these approaches, such as cell-surface located molecules, metal uptake genes and different transporters. (R-)IVET has been used already in several pathogen-related studies. Interestingly, with their studies on *L. plantarum*,the authors found considerable overlap in the expression of host-related genes with those identified in similar studies on pathogenic bacteria. A nice link to the paper of McGrath *et al.* in this issue was provided by the idea to couple IVET-identified promoters to bacteriophage lytic cassettes in intestinal bacteria, which could be used for the delivery of desirable proteins to specific sites in the gut.

The impact of bacteriophage genomics on food biotechnology is discussed in the article by McGrath, Fitzgerald and van Sinderen. The authors focus on two major aspects of bacteriophages: spoilage of food fermentations and how to avoid it, and bacteriophage-based therapies to attack pathogenic bacteria. In fact, phages have provided genomics cases 'avant la lettre' since publication of the first full phage genomic sequence (that of phage ϕX174) back in 1977. It became clear that (pro)phages play important roles in processes as diverse as horizontal gene transfer, chromosomal rearrangements, and the introduction of virulence factors into genomes of bacteria. To date, phages have played an important role in food biotechnology and many comparative analyses have been carried out, resulting in the development of rapid identification techniques. Various phage resistance mechanisms have been identified and are discussed in this paper, although it still takes a considerable effort to introduce these resistance mechanisms into a starter culture in a food-grade manner. A special application of phage genes is the use of lytic cassettes to induce *in situ* (in this case cheese) lysis of bacteria to deliver their important enzyme activities to the curd at an early stage. Many variations on this theme have been successfully developed, as described by the authors. A second major application area of phages is that of biotherapeutics. The availability of genome sequence information will aid the application and development of phages as antimicrobial agents that can effectively lyse specific pathogens during infection. Phage-encoded enzymes (e.g. murein hydrolases) can be applied to kill particular pathogens. Moreover, phage-derived integrases can be used for mammalian cell engineering, by enabling the integration of plasmids into the cell nucleus. Finally, genomics studies on prophages of commensal bacteria indicate that genes of these prophages may contibute to increased survival of the bacteria in the gastro-intestinal tract. With the wave of new phage genome information on the way, many new developments

in the areas discussed by the authors are expected in the next few years.

In their review, Nes and Johnsborg explore the potential of lactic acid bacteria (LAB) to produce antimicrobial compounds. As they point out, there is high demand for mild processing of food, and natural antimicrobials offer novel routes for preventing the outgrowth of pathogenic or spoilage bacteria in food. There are several different types of antimicrobial compounds produced by bacteria, varying from weak organic acids and other small organic molecules to, most prominently, antimicrobial peptides — also called bacteriocins. The authors discuss the biosynthesis and mode of action of the bacteriocins, which can be divided into two main classes: the post-translationally modified lantibiotics and the unmodified linear species. Interestingly, it appears that many bacteriocins interact with specific receptor molecules in the membranes of sensitive bacteria. In the case of the lantibiotic nisin, this receptor has been identified as the lipid II molecule. Genomic analyses are being used for the rapid discovery of new bacteriocins, making use of specific signatures, such as those of leader peptides, or by inspecting surrounding regions of putative structural genes for biosynthetic or immunity genes. In the pathogenic bacteria *Streptococcus pneumoniae* and *Streptococcus mutans* seven new putative bacteriocins were identified in this way. Transcriptome analysis will help to unravel the regulatory mechanisms operating during their biosynthesis under varying conditions. Another class of antimicrobials comprises low molecular weight compounds produced by specific food and gut bacteria. Some molecules belonging to this class have the interesting and highly relevant property of being antifungal; examples of such antimicrobials are cyclic dipeptides and β-hydroxypropionaldehyde. Recent analyses, employing genome mining and transcriptome analyses, have started to unravel the biosynthetic pathways and regulation of genes involved in the biosynthesis of these compounds. With the genome sequences of more than 15 LAB soon to be available, there is a wealth of information coming available on the 'antimicrobiome' (i.e. the full potential of antimicrobial substances of these organisms).

Bioinformatics is indispensable for successfully applying genomics and related strategies to understand and improve bacteria relevant for food biotechnology. Siezen and coworkers nicely overview the rapidly expanding field of microbial computational genomics, with an emphasis on genome mining and comparison of various LAB, stimulated by the fast progress in sequencing and annotation.

As various important traits of LAB are located on plasmids and conjugative transposons, efforts are also being made to obtain sequences of these entities. Various tools for the accelerated assembly, annotation and visualization of genomes are discussed by the authors. This information, together with that of publicly available data on unfinished genomes, presents key opportunities for comparative genomics. The databases and tools developed will enlarge our knowledge on, for example, diversity, evolution, horizontal gene transfer and network architectures. Visualization of the data by atlases, bar-code plots or two-dimensional projections can then be linked to other information (e.g. the ecological niches of the bacteria studied). For construction of phylogenetic trees novel bioinformatic tools are available that use combinations of conserved gene sequence information. The next challenge is to couple genomics to transcriptome analyses. Several methods for the reconstruction of gene regulatory networks have been published or are being developed, which will help us to understand the dynamic behaviour of bacteria growing under varying circumstances. This knowledge can be further extended and validated by proteome analyses. Recently, the first databases with proteome data obtained from two-dimensional gels and mass spectroscopic analyses were generated for *L. lactis*. Using programs that predict *cis* regulatory sequences in genomes helps to complete the reconstruction of pathways and gene regulatory networks. The next big challenge will be to develop bioinformatics tools and frameworks to integrate the wealth of information produced by experimental studies. Metabolome and flux analysis, kinetic modeling, interactome analysis and functional characterization of proteins are required, next to the other 'omics' approaches, to fully equip us to develop the first *in silico* models of bacterial functioning. In fact, one of the best-characterized LAB, *L. lactis*, with its small genome, relatively simple metabolism and high industrial relevance, provides an excellent case for starting a multidisciplinary systems biology effort, integrating the rapidly expanding knowledge on this microorganism. Such initiatives will catalyse and aid similar efforts on other food and gut bacteria, with the attractive perspective of being able to use cellular simulation models as knowledge repositories and additional aids in research.